RUNNING HEAD: THE EFFECT OF ENVIRONMENTAL FIELD TRIPS ON STUDENT

The Effect of Environmental Field Trips on Student Learning in Middle School

Robert S. Legutko

May 2005

Abstract

The effect of an environmental field trip on student learning in one middle school in the mid-Atlantic region of the United States was examined. An experimental one-group pretest-posttest group design was implemented on a sample of 579 students which comprised 12 groups. Although a *t*-test for dependent samples indicated that less than half of the groups showed a statistically significant improvement from pretest to posttest at the $p < .05$ level, a group item analysis indicated that in only 1 out of 60 instances was the group mean lower than 70% on a posttest question (and in 36 out of 60 instances, the group mean was 90% or higher on a posttest question), indicating that most students were able to demonstrate learning of the material contained within the field trip immediately afterwards.

Introduction

This study constituted the first step in developing a research effort to evaluate the benefits of environmental field trips in one state in the mid-Atlantic region of the United States. The focus was specifically on the funding of one program in a particular county which allows public and parochial school children to experience designated educational destinations that meet this state's primary and secondary education academic standards.

The basic research question that guided this study was, "What effect do environmental field trips have on student learning?". A secondary question, "What effect do environmental field trips have on student enthusiasm for learning?" was originally proposed but was not addressed in this study. A qualitative assessment to measure enthusiasm, attentiveness, time on task, and energy levels of student participants is recommended for future study.

Purpose of the Study and Problem Statement

The purpose of the study was to determine the educational value and efficacy of environmental field trips on student learning. The problem is that it was unknown whether or not students were learning (or sufficiently learning) on the field trips currently undertaken. If sufficient learning occurred,

then the field trips achieved their intended goal(s), and it could be argued that they were worth the time, money, and effort put forth by all parties involved. If sufficient learning did not occur, it could then be argued that the resources either might instead be allocated in more beneficial ways or eliminated altogether.

Participating Venue and Students Involved

There are currently 10 participating venues from which the public and parochial schools may choose field trip participation in the county's funding program. The venue utilized in this study was a non-profit, member-supported organization dedicated to land preservation, river restoration, trail development and environmental stewardship through education for over 30 years.

Sixth-grade students from two suburban public middle schools were utilized in this study. The district educates approximately 8,500 students, and the median annual household income is approximately $65,000. The two middle schools combine to educate approximately 2,100 students, with an ethnic composition of 91% Caucasian, 5% Asian, 2.5% Hispanic, and 2% African-American students.

Review of Related Literature

Unfortunately, many administrators balk at requests for field trips due to increased liability concerns and the movement towards academic accountability. The focus on scoring well on standardized tests causes them to question the value of field trips (Myers & Jones, 2003).

But one obvious positive effect of field trips is that students enjoy getting out of the classroom and having real-world, firsthand experiences. Flexer and Borun (1984) examined fifth and sixth grade students at the Franklin Institute Science Museum in Philadelphia and found that the strength of the science museum exhibit was in the affective domain. Students found the exhibit much more enjoyable, interesting, and motivational than a classroom lesson. Interestingly, however, these students scored lower on a content assessment than a comparison group attending a classroom lesson in the museum on the same content. Knapp (2002) also discovered that posttests (at 1-month and 18-month intervals) of elementary school students who had participated in an environmental science program at a community park in a Midwestern city yielded both positive responses toward wanting to learn more about the subject matter and an interest in returning to the field trip site, but students also had nonspecific and disassociated memories from information given by the field teacher. This meant

that while not retaining specific objective-oriented content, students did gain a positive reaction to returning that continued over the long-term.

A cultural benefit for otherwise culturally-deprived students was realized in a 2002 study by Floyd. Approximately 50 at-risk sixth-grade African-American students from three inner-city South Carolina middle schools participated in a week-long summer program at the Columbia Museum of Art to help them understand artists' lives and art works and to further understand their own worlds. Floyd reported that the students walked away from the museum feeling, among other things, as if they actually knew the collector.

Teachers also have an important role in field trips. Myers and Jones (2003) indicated that they should be actively engaged in teaching activities throughout the field trip and utilize different teaching approaches than in traditional classrooms, such as interacting with students to help answer questions, initiating discussion with small groups of students by asking them questions, and functioning more as facilitators or guides. They further stated that teachers can also increase student interest and learning by playing an active rather than a passive role during the field trip.

Also regarding teachers' roles, it has been found that providing follow-up and culminating activities are essential for

the reinforcement of student learning. A culminating activity of some kind should give the students an opportunity to apply the content knowledge they gained during the field trip, help students tie together both content they covered in regular class sessions and content learned during the field trip. These can be whole class or small group experiences, and should occur as soon after the trip as possible (Myers & Jones, 2003). This effect was demonstrated in a study by Farmer and Wott (1995) with 111 fourth grade public school students who were all enrolled in the same school and had completed similar course work related to the topic on a field trip to Washington Park Arboretum in Seattle. A pretest-posttest control group design found that there were significant differences between all pretest and posttest scores, primarily because the related follow-up activity reinforced some of the concepts presented during the field trip.

Methodology

The design implemented in this study was an experimental one-group pretest-posttest group design to determine whether or not the independent variable (environmental field trip) had a significant effect on the dependent variable (student learning). Specifically, the design helped determine whether or not the number of students who answered the posttest items correctly,

compared to the pretest, was higher to a statistically significant (i.e., meaningful) degree.

The threats to internal validity in this design included history, maturation, mortality, instrumentation, testing, and statistical regression (Gay, Mills, & Airasian, 2006). Three of these threats -- history, maturation, and mortality -- were not factors because of the short one-day time frame of this study. Instrumentation was also not a threat to internal validity because the exact same assessment was administered in the exact same manner as both the pretest and posttest. Testing may have been a threat because of pretest sensitization and the short one-day time frame between pretest and posttest. Statistical regression, the tendency for the posttest scores to gravitate towards the mean, was not controlled in this study.

The threat to external validity in this design was pretest-treatment interaction, which may have been present because some participants could have recalled test questions from the pretest to the posttest due to the short one-day time frame of this study. Borg, Gall, and Gall (1993) stated that research findings are “externally valid to the degree that their results can be generalized to persons, settings, and times different from those involved in the research” (p. 303). No generalizations should be made from this study for populations that are demographically different from the participants involved.

Values for *p* were calculated by adding number of correct responses for each group on all pretest and posttest questions, then analyzed using *t*-test for dependent samples at the $p < .05$ level.

Procedures

The initial evaluation was piloted in October 2004 on 6th grade students from the same school district. The pretest and posttest questions were in multiple choice format and administered in written form. Based upon both the evaluation results and suggestions by the venue site administrator, it was determined that data would instead best be gathered by orally questioning students with multiple choice items in groups because students may have been intimidated by the formality of a written pretest and posttest. It was also noted that it could not be guaranteed that regular classroom teachers would consistently (if at all) administer pretests and posttests off-site.

In May and June 2005, 12 groups of students (579 students overall) from the two middle schools participated in field trips at the venue site. Students were informed upon arrival that they would be surveyed as a group. For the pretest, students were instructed to cover their eyes and raise their hands to respond to each question. The hand count was then recorded by the venue

site administrator. Students then participated in the field trip (approximately three hours). After the field trip, the exact same covering eyes and raising hands questioning procedure was followed for posttest items. The responses were then collected and analyzed using a *t*-test for dependent samples. A copy of the pretest/posttest items is found in Appendix A.

Results

Data were analyzed individually for each of the 12 groups in this study by adding the correct responses for each Item Number (1 through 5) on both the pretest and posttest. For example, Group 1 had 32 correct pretest responses and 53 correct posttest responses on Item Number 1, 20 and 51 on Item Number 2, 14 and 55 on Item Number 3, 33 and 54 on Item Number 4, and 33 and 57 on Item Number 5. Using these five groups of numbers in a *t*-test for dependent samples, the improvement from pretest to posttest was statistically significant for Group 1 at the $p < .0001$ level. Data for all 12 groups is listed in Table I. Altogether there was a statistically significant difference in group improvement at the $p < .05$ level or better in only 5 out of the 12 groups.

The percentage correct on the posttest for each Item Number is also listed in Table I. For the 60 possible Item Number responses on the posttest (5 Item Numbers multiplied by 12

Table I

Number and Percentage Difference between Pretest and Posttest Scores by Group and Item Number

	Item Number 1			Item Number 2			Item Number 3			Item Number 4			Item Number 5			*p*
	pretest	posttest	difference	pretest	posttest	difference	pretest	posttest	difference	pretest	posttest	difference	pretest	posttest	difference	
Group 1																
n correct	32	53	21	20	51	31	14	55	41	33	54	21	33	57	24	0.0001***
% correct	56%	93%	37%	35%	89%	54%	25%	96%	72%	58%	95%	37%	58%	100%	42%	
Group 2																
n correct	40	40	0	34	38	4	9	40	31	39	43	4	39	45	6	0.1735
% correct	87%	87%	0%	74%	83%	9%	20%	87%	67%	85%	93%	9%	85%	98%	13%	
Group 3																
n correct	34	42	8	26	33	7	9	43	34	22	41	19	31	45	14	0.0094**
% correct	72%	89%	17%	55%	70%	15%	19%	91%	72%	47%	87%	40%	66%	96%	30%	
Group 4																
n correct	29	35	6	18	35	17	26	32	6	30	37	7	31	37	6	0.0105*
% correct	63%	76%	13%	39%	76%	37%	57%	70%	13%	65%	80%	15%	67%	80%	13%	
Group 5																
n correct	40	43	3	13	28	15	24	44	20	33	37	4	27	45	18	0.0616
% correct	87%	93%	7%	28%	61%	33%	52%	96%	43%	72%	80%	9%	59%	98%	39%	
Group 6																
n correct	41	43	2	15	40	25	2	39	37	36	39	3	35	42	7	0.0829
% correct	93%	98%	5%	34%	91%	57%	5%	89%	84%	82%	89%	7%	80%	95%	16%	
Group 7																
n correct	45	47	2	11	42	31	33	44	11	30	40	10	42	48	6	0.0874
% correct	90%	94%	4%	22%	84%	62%	66%	88%	22%	60%	80%	20%	84%	96%	12%	
Group 8																
n correct	48	49	1	19	41	22	30	43	13	20	36	16	43	48	5	0.1105
% correct	98%	100%	2%	39%	84%	45%	61%	88%	27%	41%	73%	33%	88%	98%	10%	
Group 9																
n correct	51	50	-1	8	44	36	20	48	28	23	47	24	48	50	2	0.0678
% correct	100%	98%	-2%	16%	86%	71%	39%	94%	55%	45%	92%	47%	94%	98%	4%	
Group 10																
n correct	41	44	3	8	40	32	19	42	23	20	44	24	27	44	17	0.0069**
% correct	91%	98%	7%	18%	89%	71%	42%	93%	51%	44%	98%	53%	60%	98%	38%	
Group 11																
n correct	46	48	2	6	45	39	37	48	11	41	46	5	46	48	2	0.1552
% correct	92%	96%	4%	12%	90%	78%	74%	96%	22%	82%	92%	10%	92%	96%	4%	
Group 12																
n correct	34	47	13	21	43	22	20	47	27	36	47	11	32	46	14	0.0010**
% correct	71%	98%	27%	44%	90%	46%	42%	98%	56%	75%	98%	23%	67%	96%	29%	

Note. Data for columns labeled pretest and posttest indicate the number and percentage of students in each group who responded correctly to the item number indicated. Data for columns labeled difference indicates cumulative and percentage improvement for each group on the item number indicated. Values for *p* calculated by taking the sum of correct responses for each group on pretest item numbers 1-5 and posttest item numbers 1-5, then analyzed using *t*-test for dependent samples.

*$p < .05$; **$p < .01$; ***$p < .001$

groups), groups were correct from 70%-79% on 5 posttest items; from 80%-89% on 18 posttest items; and from 90%-100% on 36 posttest items. Cumulatively, the groups scored at least a 70% on 59 out of 60 posttest items (for a 98% success rate).

A four-point Likert-scale questionnaire that sought attitudes about the field trip to this site (which contained similar, but not identical content) was given to the 44 students that comprised the October 2004 pilot group. There was a mean of 3.36 for the question, “How much did you like (the field trip site)?”, a mean of 3.52 for the question, “Would you want to go on another field trip like the one at (the field trip site)?”, and a mean of 3.30 for the question, “How much do you think you learned from the field trip?”. This indicated that students responded positively in the affective domain regarding their field trip visit to this venue.

Discussion

Less than half (5 out of 12 groups) showed a statistically significant improvement from pretest to posttest. This result was not anticipated (it was anticipated to be higher), especially since the questions asked were identical in the pretest and posttest, and because the field trip itself was only about three hours in length. Forecasting of the questions in the pretest should have served as an anticipatory set for the

students before the field trip, and student consequently should have been able to correctly respond to the same questions in the posttest activity. This did not occur as frequently as expected. Statistically speaking, then, the educational value of field trips in this instance would have to be in question.

It should be noted, however, that in only 1 out of 60 instances was the group mean lower than 70% on a posttest question (and in 36 out of 60 instances, the group mean was 90% or higher on a posttest question). This would indicate that most students were able to demonstrate learning of the material contained within the field trip immediately afterwards.

It should also be noted that having students experience real-world activities outside of the classroom may have positive results not measured in this or any similar short-term study. Student responses from the October 2004 pilot group overwhelmingly indicated in their written questionnaire that they specifically liked the field trip – and not necessarily that they liked getting out of the class.

Limitations and Recommendations

Rudman (1994) stated that there are weaknesses in field trip research designs. Studies often do not define which dimensions of the cognitive domain are being measured, little

data is found on the longevity or retention of learning, and there is little consideration for special needs visitors.

In this design, the sample size was relatively small (579 students) and narrow (only sixth-graders, 91% Caucasian, affluent suburban school district) and may not translate well to the population.

Another limitation of the study could be remedied by examining two equal groups of students, teaching content to one group solely in the classroom and by classroom methods, and teaching the same content to the second group solely through the field trip. Assessing and comparing the groups would indicate a more direct comparison of classroom methods versus field experiences.

A study to determine the effect of environmental field trips on student enthusiasm for learning would also be recommended as a future study, involving a qualitative assessment by venue staff and/or teachers measuring enthusiasm, attentiveness, time on task, and energy levels of the student participants.

Conclusions

It may be concluded that significant learning was not achieved as frequently as anticipated or desired by the student participants on their field trips to the venue in this study.

The educational value from a purely pedagogical standpoint is questioned. However, the practical experiences – bringing bookwork to life – may not be entirely measurable, and may, in fact, be the entire point and purpose of field trips in general. Can “number crunching” truly measure, or should it measure, the affective value of student learning? Because a student does not respond correctly on the test, does it mean that he/she truly did not learn? Perhaps educators need to reexamine exactly what it is that students should be learning in its hidden curriculum, the behaviors, attitudes, and knowledge that the school unintentionally teaches its students (Parkay & Stanford, 2002) and continue to provide those types of life-building experiences.

References

Borg, W. R., Gall, J. P., & Gall, M. D. (1993). *Applying educational research: A practical guide* (3rd ed.). New York: Longman.

Farmer, A. J., & Wott, J. A. (1995). Field trips and follow-up activities: Fourth graders in a public garden. *Journal of Environmental Education, 27*(1), 33-35.

Flexer, B., & Borun, M. (1984). The impact of a class visit to a participatory science museum exhibit and a classroom science lesson. *Journal of Research in Science Teaching, 21*(9), 863-873.

Floyd, M. (2002). More than just a field trip: Making relevant curricular connections through museum experiences. *Art Education, 55*(5), 39-45.

Gay, L. R., Mills, G. E., & Airasian, P. (2006). *Educational research: Competencies for analysis and application* (8th ed.). Upper Saddle River, New Jersey: Merrill Prentice Hall.

Knapp, D. (2002). Memorable experiences of a science field trip. *School Science and Mathematics, 100*(2), 65-72.

Myers, B., & Jones, L. (2003). Successful field trips: A three-step approach. *The Agricultural Education Magazine, 75*(4), 26-27.

Parkay, F. W., & Stanford, B. H. (2004). *Becoming a teacher* (6th ed.). Boston: Allyn and Bacon.

Rudmann, C. L. (1994). A review of the use and Implementation of science field trips. *School Science & Mathematics, 94*(3), 138-141.

Appendix A

Pretest and Posttest Items

1. What do we call a community of living organisms and their interrelated physical and chemical environment?
   a. a population
   *b. an ecosystem
   c. a tribunal

2. What will environmental degradation do to the diversity of an aquatic community?
   a. eliminate tolerant organisms
   *b. eliminate intolerant organisms
   c. make no difference

3. What kind of effect does a well-established riparian zone have on stream health?
   *a. a positive effect
   b. a negative effect
   c. no effect

4. Will developing a riparian zone impact stream health if lawns -- instead of parking lots -- border the waterway?
   *a. yes
   b. no

5. Are all of our local owls are at the top of the food chain?
   a. yes
   *b. no

(correct responses designated by asterisk)